STRESS DETECTION FROM CHAT USING NATURAL LANGUAGE PROCESSING EMBEDDINGS AND TOKENIZATION USING OCIR MACHINE LEARNING METHOD

[1]Dr.S.Brindha, [2]Dr.S.Ravichandran, [3] Sajith Balaji Sarvesh

[1] Assistant Professor, Department of Computer Applications,
College of Sciences, SRMIST Faculty of Science and Humanities,
SRM Institute of Science and Technology, SRM University, Kattankulathur, Chennai, Tamilnadu.
[2] Assistant Professor and Head, Department of Chemistry, St.Peter's Institute of Higher Education and Research, Chennai, Tamilnadu, India.
[3]Student, Department of Computer Science and Engineering, Karpaga Vinayaga College of Engineering and Technology, Chengalpattu, Chennai, India
Corresponding Author
Dr.S.Brindha
brindhas1@srmist.edu.in

ABSTRACT

Stress is a subjective sensation that is difficult to accurately define. Stress can have a variety of biological and psychological consequences on one's health, which can be defined and quantified. We only get half of the information if we focus solely on what someone says and ignore what their demeanor tells us. In this project, we will use random forest and decision tree algorithms to detect mental stress. The existing system does not work in real time and is inaccurate and inefficient in terms of loading and implementation durations. Furthermore, the appropriate test-train split ratio is not applied during testing and training. The OCIR proposed method is implemented in real time and has a high level of accuracy. In comparison to the current system, the suggested solution has incredibly fast loading and execution times. The OCIR proposed method can be improved for more complex use cases and is highly effective and scalable.

Keywords: Machine Learning, Advanced Strees Detection from Chat, Natural Language Processing, Tokenization, Text Conversation.

CHAPTER 1

INTRODUCTION

With the rapid expansion of society and the economy, an increasing number of people lead stressful lives. Too much stress jeopardizes both physical and psychological health. Because of their spiritual immaturity, the threat is particularly dangerous for young people, who are more likely to experience sadness or even suicide. As a result, psychologists and educators have focused heavily on the topic of adolescent stress [3]. However, one major issue is that most developing adolescents are unwilling or hesitant to disclose their sentiments to others, preferring to relieve tension in the virtual world. Stress is a sensation of mental or physical tension caused by a person's inability to cope with the surroundings. Stress detection is a classification task that determines whether a specific target is stressed. The task has piqued researchers' interest for two reasons: first, stress detection is important in applications such as psychological well-being, cognitive behavior therapies, and safe driving; second, stress is a known regulator of human emotion mechanisms, so stress detection research may benefit the development of emotionally intelligent agents. Stress has a variety of effects on human behavior. A person's standard of living is greatly influenced by his emotional states, such as stress and anxiety [1].

Stress is often described as a complex psychological and behavioral state resulting from the perception of a significant imbalance between the demands placed on the individual and their perceived ability to meet the demands[10] . Stress has an impact on both mental and physical health, generating disorders such as irregular heart rhythms, arrhythmia, and depression.

CHAPTER II
LITERATURE REVIEW

Jindal et al. 2019 offer a work that uses Natural Language Processing (NLP) approaches to detect stress in online chat chats. The authors present an approach for preparing chat data, extracting features, and classifying them using machine learning methods. The study employs a dataset of real-world chat chats to evaluate the effectiveness of various classifiers, including SVM, KNN, and Random Forest. The findings indicate that the proposed approach is successful in detecting stress in online chat chats, with an accuracy of up to 87%. This study has significant implications for creating automated methods to detect stress in online conversation and giving prompt assistance to people in need[6].

Buechel, S., et al. 2019 define this paper as a study on identifying emotional distress and stress in social media messages using natural language processing. The authors suggest a way for constructing a multi-task learning framework that models both stress and distress classification. The method employs a dataset of Twitter messages to examine the performance of several models, including neural network models and standard machine learning classifiers. The results reveal that the suggested method outperforms other methods for detecting stress and emotional distress in social media messages [7]. This study has significant implications for better understanding mental health concerns and designing tailored interventions to help people in distress.

Lee, C. H., et al. describe that this paper presents a chatbot system for stress management based on deep neural networks. The authors describe a framework for creating a chatbot that can recognize users' emotional states and offer suitable stress management strategies. A dataset of stress-related terms and phrases is used to train a deep neural network model that can classify users' emotional states [8]. The chatbot system recommends stress management measures based on the user's emotional state and initiates a discussion with the user to offer emotional support. The study assesses the proposed system using user surveys and finds that the chatbot is successful at delivering stress management support to users. This study has significant implications for the development of automated systems for mental health support.

Kudugunta, S., et al. demonstrate a method for identifying stress in email exchanges using natural language processing (NLP) approaches. The authors propose creating a stress detection algorithm by extracting linguistic and discourse elements from email discussions. The study employs a dataset of real-world email interactions to assess the efficacy of several machine learning classifiers, including SVM, Logistic Regression, and Random Forest. The findings indicate that the proposed approach is successful in detecting stress in email exchanges, with an accuracy of up to 78%. This study has significant implications for building automated methods to detect stress in email conversation and offering prompt assistance to persons in need [10].

CHAPTER-III
PROPOSED WORK

Incoming photos are categorized into distinct classes, and suitable preprocessing is conducted for each group. The resultant images serve as input for the OCIR engine. Additionally, certain post-processing techniques may be implemented on the output of the OCIR engine. A dataset of photographs from a social network has been compiled, and experiments have been conducted using this dataset.

3.1 Pre-processing

With the goal of returning only the most relevant pages, "sparse search methods" attempt to reduce the dimensionality of the indexed content. These methods work on the assumption that getting all of the documents from a collection would be pointless as only a small subset of the data is pertinent to a given query.

Data Collection from Text: An analysis of the incoming social network image stream was conducted, and a representative group of images primarily in Russian containing potential social propaganda and information leakage were selected. The 67 photos that make up the input dataset include 64554 symbols. Four distinct groups have been established for all of the photographs.

Demotivators. Social networks frequently employ photographs of this type. Typically, each photograph has a main image (scene, people, etc.) and some text underneath it. Twenty-four percent of photographs are in this category.

Photographs of documents, certificates, identity cards, etc. are typically included in this category. Some sentences of text may be included in images in this category. There are 38% pictures in this category. Examined the Images in this category are high-quality scans of papers that include a lot of text. Twenty-four percent of photographs are in this category.

Smartphone. The photographs in this category are the same as those from the scanned category, but they were taken using a low-quality smartphone camera. There are 14% of photos in this category. Mostly 85% of the portrait galaxy essential remain engaged by the text document. Images from certificates and demotivators may include a background, a scene, and typically a few phrases of text. On the other hand, photos from scanned categories and smartphones typically contain a lot of text and are a picture of a text document.

3.2 Feature extraction

Opinion mining is increasingly recognized as a burgeoning area of research. The objective is to align specified textual sequences by a technique known as pattern matching. Preparation for demotivators. Demotivator class graphics generally use a dark background with light text positioned below the primary image. Dark text may subsequently manifest on the primary image. Our experiments, however, demonstrated that Tesseract exhibits markedly superior performance when the text is dark against a light background. Moreover, certain demotivators may feature text with markedly inconsistent font sizes, thereby diminishing identification accuracy.

Consequently, we employ three iterations to process images of this category. Upon completion of the requisite preprocessing, the image is supplied to Tesseract, which receives the text and corresponding bounding boxes at each stage. Text segments recognized at each phase are subsequently amalgamated.

All extensive Hough lines of the image are encompassed inside the minimal rectangle that delineates the primary image. During the second iteration, the complete image undergoes bitwise inversion, and the background color is superimposed onto the primary image. All bounding boxes from previous iterations are obscured with background color and input into the OCIR engine during the third iteration. Preparation of certificates. Images in this category may exhibit low resolution, and a small segment of the image may contain text. Furthermore, the presence of background elements in the primary document input image diminishes recognition accuracy. Consequently, we utilize the preprocessing outlined below.

Improvement of image resolution: It is applied to images that are below a specified threshold size. The principal document is subsequently extracted from the input image through the document localization procedure. The background color is selected from the identified bounding boxes, while the rest of the image is obscured. The preprocessing has been scanned. Text recognition algorithms perform effectively with photographs of this category. Resolution enhancement is implemented solely when the image's area falls below the threshold. In the initial phase of preprocessing for smartphones, lighting improvement is applied to the input image if deemed essential. When the image's area falls below the threshold, an enhancement in image resolution is implemented. The subsequent phase entails identifying perspective distortions and, upon their detection, implementing the appropriate perspective correction to the image.

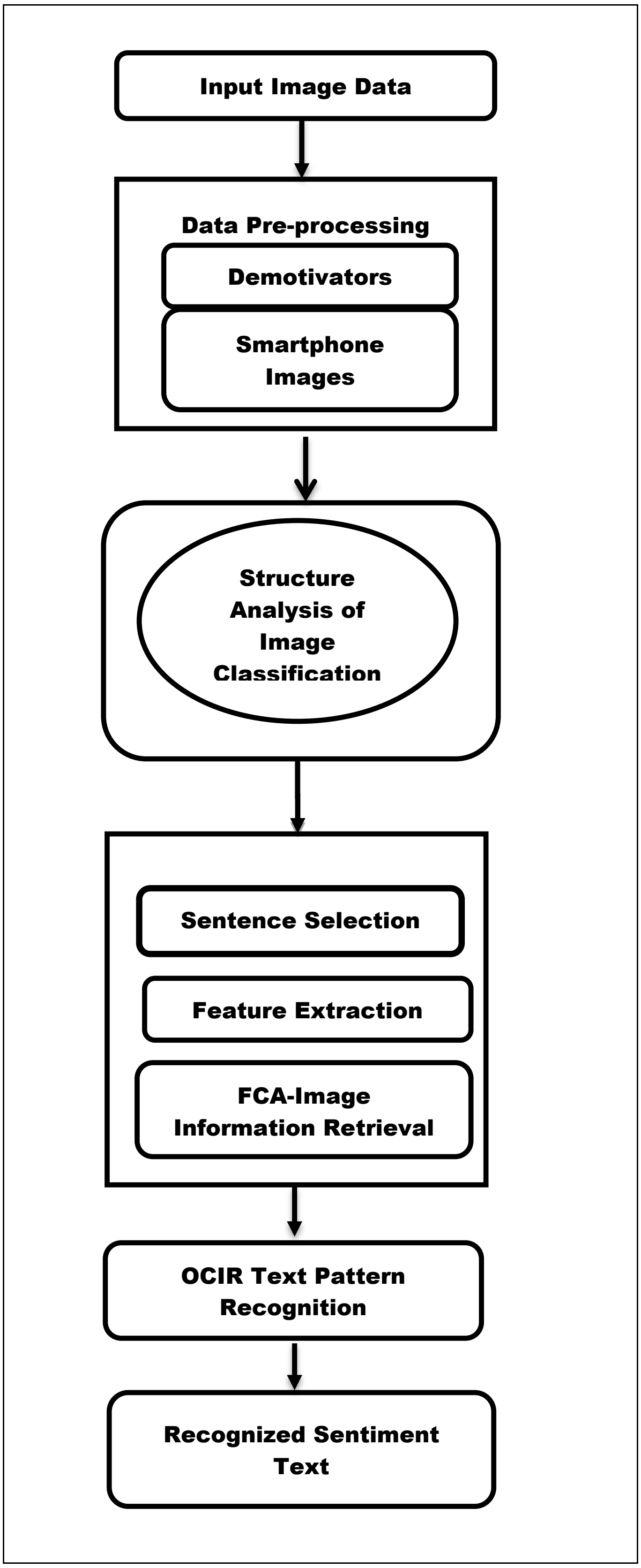

Fig. 3.1 Flow Diagram of OCIR

The suggested pipeline facilitates the extraction of textual information from social media, when a substantial volume of data is conveyed through low-quality photos or intermixed with non-textual content. The OCIR engine utilizes the input images to generate the output text. If the input image corresponds to the demotivator class, supplementary rounds of the second and third stages are implemented. Words exhibiting low confidence are excluded from the final text if the overall confidence of the document falls below a specified level (an extra parameter). Predefined classes are utilized to classify input photographs. Subsequently, the vibrant image is converted to grayscale and undergoes class-specific preprocessing. Additional preprocessing and recognition iterations are required for the "demotivator" category. The final text file is generated following the post-processing of the Tesseract engine's output to eliminate certain extraneous symbols.

3.3 Fuzzy Concept Analysis for Information Retrieval Method (FCA-IR)

By concentrating on the allocation of specific classes instead of categories throughout the document's trajectory, text classification enhances its organization in a coherent and user-friendly format. In the supervised method, the model is developed using the training set. The categories are predefined, and the documents in the preparation dataset are manually annotated with one or more category labels. Upon training on the earnings dataset, the classifier can subsequently predict the category of a new document. The classifier's strength, contingent upon the employed classification algorithm or approach, yields a confidence estimate reflecting the certainty of the organization's label's accuracy.

The impact of the stipulations in the document. Any statement possessing semantic significance is classified as a term, and the document's meaning is augmented by the aggregation of these words. Terminological methods address the contemporary issues of polysemy and synonymy. A term with several meanings is termed polysemic; a term with various meanings is referred to as synonymous. The semantic interpretations of numerous newly identified terms are vague and do not meet user expectations. The two principal categories of text mining techniques are phrase-based and term-based methods.

3.4 Stemming

Stemming is the procedure of eliminating derivational and inflectional affixes from words to revert them to their base form. Stemming is frequently employed in occupations related to information retrieval.

Step 1:	***Input image***
Step 2:	***If (Wn == 0)*** ***Return final Sentiment (F_p, e_P)***
Step 3:	***Else if *Wp== 0)*** ***Return final Sentiment (F_n, E_n)***
Step 4:	***Else*** ***{*** ***If (Fp- Fn >0.1)*** ***Return final sentiment (F_r, E_r)}***
Step 5:	***If (Fp + Fn >0)*** ***Return final sentiment (F_r, e_r)***
Step 6:	***converted plain image to text considered to be a meaningful English sentence but it captures the original sentiment and context.***
Step 7:	***Sentiment polarity classification system M is a quadruple*** ***$M=\{\alpha,\lambda,\delta,\varphi\}$***
Step 8:	***Function to calculate the similarity score of two messages.***
Step 9:	***Using accuracy, determine the k nearest neighbours that are the -***

	heuristically optimal number.
Step 10:	***Assess the reaction to text messaging on social media and other web channels.***

Numerous research indicate that stemming enhances the efficacy of information retrieval systems. The Porter stemmer is the preeminent algorithm for English stemming. Nevertheless, this stemming technique possesses numerous limitations as English morphology cannot be adequately elucidated by its basic principles. Stemming is the process of reducing an inflected term to its root form. The names "programming," "programmer," and "programs" can all be abbreviated to the shared root "program." In other terms, "program" is synonymous with the three preceding inflection words. This text-processing method is frequently advantageous for mitigating sparsity and/or standardizing terminology. The similarity in meaning between inflected words and their stems minimizes redundancy and enables natural language processing (NLP) models to discern linkages between them, hence enhancing the model's comprehension of their usage in related settings.

Fig.3.2 Text and Image Extraction

From 3.2 text and image extraction using the stemming algorithms truncate the beginning or end of a word utilizing a compilation of common prefixes and suffixes present in inflected terms. We can ascertain that while this approach offers advantages, it also presents drawbacks, as it sometimes produces word stems that are not legitimate words. This technique yields optimal model parameters, including network design, learning rate, and regularization, which enhance the accuracy of blood group prediction. The output results indicate that the integration of AI, image processing, and FA significantly enhances both the accuracy and efficiency of Bgroup categorization.

3.5 Stop Words

Stop words are integral components of the lexicon commonly utilized across all languages, not exclusively English. Everyday vocabulary in any language, not alone English, inherently encompasses. Terms frequently employed across various languages, not exclusively English, are referred to as stop phrases.

The stop phrase incentive is effective for several items as it enables the emphasis on distinctive terms while eliminating relatively common ones in a comprehensible manner. Many individuals perceive stop words as a singular collection of phrases. It can undoubtedly present a diverse array of subjects in multiple circumstances. Examples of coordinating conjunctions that unite words, phrases, and clauses include for, and, nor, yet, or, but, and so. The temporal significances of prepositions are disparate.

CHAPTER 4

SYSTEM ANALYSIS

4.1 CHARACTERISTICS OF EXISTING SYSTEM

The Support Vector Machine (SVM) technique is the foundation of one current approach for stress detection from conversation that uses tokenization and natural language processing (NLP) embeddings. SVM is a well-known machine learning technique that may be applied to classification applications, such identifying tension in chat data. The limited interpretability of SVM is another drawback. It can be difficult to pinpoint the precise characteristics that influence stress levels in chat data since the hyperplane produced by SVM is frequently hard to understand. Lastly, because stress levels may not be distributed equally across conversation participants, SVM may be sensitive to imbalanced datasets, which can be troublesome when working with chat data. Overall, SVM is a well-liked technique for stress identification from conversation utilizing tokenization and NLP embeddings, but it has several drawbacks that may restrict its usefulness in particular circumstances. When creating stress detection systems, other machine learning algorithms like Random Forest and Decision Trees should be taken into account since they may have certain advantages over SVM.

4.2 CHARACTERISTICS OF PROPOSED SYSTEM

In order to find patterns in chat data linked to stress, the suggested system for stress detection from chat utilizing natural language processing (NLP) embeddings and tokenization will make use of a number of NLP approaches. The capacity of NLP to handle unstructured text data and get useful information from it is one of its key benefits. NLP methods like sentiment analysis, topic modeling, and named entity identification can be used to find patterns and trends in chat data that are suggestive of stress in the context of stress detection from chat. All things considered, the suggested approach for stress detection from chat employing natural language processing embeddings and tokenization has a number of benefits that make it a successful method for detecting stress in chat data. Researchers and practitioners can learn more about chat participants' emotional states and pinpoint conversational topics linked to stress by utilizing NLP approaches. This can enhance people's general wellbeing and help them better control their stress levels.

4.3 SYSTEM DESIGN

SYSTEM ARCHITECTURE

All of the elements that are now incorporated into the system are succinctly and clearly described in this diagram. The figure illustrates the connections between the various decisions and activities. One may argue that the entire procedure and its execution are a picture. The functional relationships between different entities are depicted in the image below.

A complete end-to-end pipeline for stress detection from text using machine learning is depicted in this diagram. The pipeline begins with the collection of raw data and ends with the prediction of the final result. Beginning with the download of text data from Kaggle, which is the source of the dataset, the process begins. In addition to this, stopwords are collected from the NLTK library. Stopwords are words that are frequently used (such as "is," "the," and "and") but do not contribute a significant amount of meaning when being analyzed. A dataframe is then created, and undesirable columns that are not related to the task are removed from the dataframe. This completes the process of organizing the dataset. The step of text preprocessing is quite comprehensive, consisting of the removal of URLs, emojis, plurals, and extraneous characters. This is followed by the removal of stopwords, which eliminates noise, and stemming, which reduces words to their root form (for example, "running" → "run") in order to standardize the text. For the purpose of visualizing frequently recurring words and assisting in the comprehension of prevalent patterns within the dataset, a word cloud is constructed using matplotlib after the cleaning process has been completed.

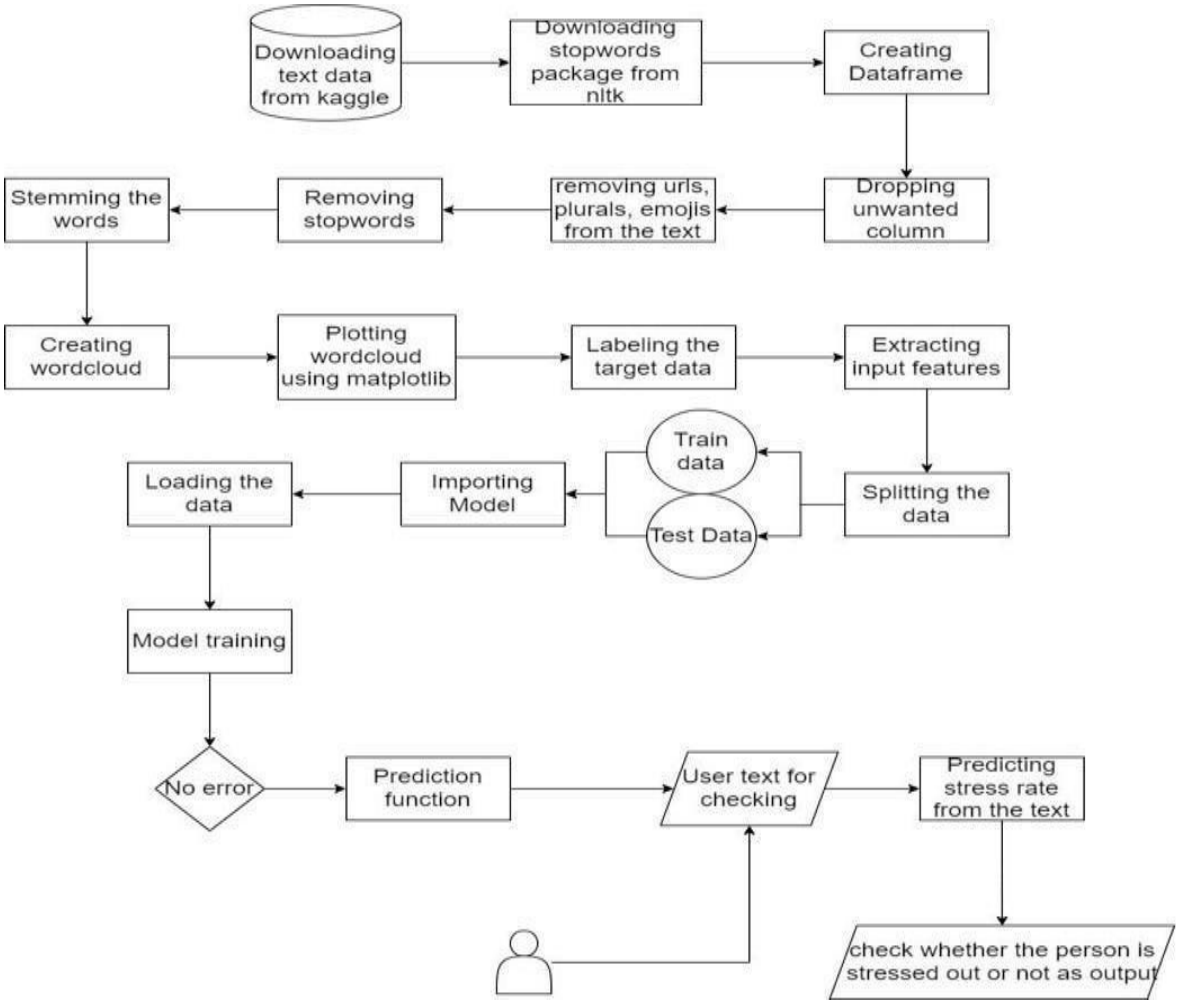

Fig 4.1 System Architecture

After that, the system moves on to the next step, which is labeling the target data. At this point, each text input is given a class, such as "stressed" or "not stressed." After that, the process of feature extraction is carried out, which often involves translating the text into a numerical format (such as Bag of Words or TF-IDF) so that machine learning models can effectively handle it. In order to ensure that the model can be tested on data that it has not before encountered, the dataset is then divided into training and testing sets. For the purpose of validation, the test data is utilized, while the model, which is most likely Naïve Bayes or a comparable model, is imported and trained using the training data. The development of a prediction function follows the completion of successful training and the verification that there are no errors.

CHAPTER 5 RESULT AND CONCLUSIONS

5.1 RESULTS

The stop phrase incentive is effective for several items as it enables the emphasis on distinctive terms while eliminating relatively common ones in a comprehensible manner. Many individuals perceive stop words as a singular collection of phrases. It can undoubtedly present a diverse array of subjects in multiple circumstances. Examples of coordinating conjunctions that unite words, phrases, and clauses include for, and, nor, yet, or, but, and so. The temporal significances of prepositions are disparate.

Table 5.1 Social Media Sentiment Analysis

Social Media Conversation	**Sentiment**	**Text Images**
Good Morning	***Positive***	
Spring Month	***Positive***	
Music	***Positive***	

Every cluster of objects undergoes pattern mining, which is dependent on the transformation process. Two potential transformations weighted and Boolean are examined. When generating a set of transaction databases {D1, D2,... D k } in both transformations, k are clustered so that the i-th transaction database D i ($1 \leq i < k$) corresponds to the i-th cluster G i. A collection of transactions {D i 1, D i2...D I |G i | } is contained in a transactional database D i, where the j -th transaction Di, i denotes the j -th item ij of the cluster G i.

Table 5.2 Social Media Text OCIR Sentiment Analysis

"I am so stressed out right now, I don't know how to handle all the work pressure."	Positive stress
"I feel like everything is going wrong and I can't control anything."	Negative stress
"I have a lot on my plate, but I am managing well and feeling confident."	Neutral stress
"I am feeling a bit overwhelmed with my new job, but I am learning and growing every day."	Positive stress
"I am so anxious about this upcoming presentation, I can't sleep at night."	Negative stress
"I am feeling a bit stressed out about my finances, but I am working on a plan to fix it."	Neutral stress
"I am feeling really burnt out from my job, I need a break."	Negative stress

Text mining and statistical analysis frequently employ TF-IDF as a weighting mechanism. The expense of tf-idf escalates in direct correlation to the frequency of a phrase's occurrences inside the collection, excluding counter-occurrences resulting from the phrase's prevalence in the corpus. This may facilitate the organization of data, resulting in specific terms appearing more frequently in files. TF-IDF is an effective pre-word filtering technique for text categorization and summarization across several problem domains. The TF-IDF object utilizes two primary metrics: term frequency and inverse document frequency. Due to the rapid proliferation of emoticons on social media, where they can enhance or condense textual meaning, researchers have examined their semantic correlation with words. Optical Character and Image Recognition (OCIR) entails converting photographs that include text into a machine-readable format. This procedure use algorithms and software to analyze photos, recognize characters, and transform them into editable text. The fundamental functionality depends on OCIR technology, which can be integrated into other software applications and internet converters.

Data mining includes several technologies for pre-processing, analysis, and interpretation. These strategies primarily categorize into two types: pattern recognition and machine learning. The objective of pattern recognition is to identify discernible evidence of verified entities and relationships, or to detect patterns within the incoming data. These processes are mostly linked to image analysis, although it is not the principal application type. Most machine learning techniques focus on deriving generalized insights from data, including images, which will then be utilized for predictive challenges. Relevant papers are those that assist the user in identifying the solution to an inquiry.

$$Character\_accuracy = \frac{Number\ of\ Characters\ errors}{Number\ of\ Characters} \quad \text{-----}(1)$$

As in our dataset classes are not balanced two kinds of accuracy are used. In the Table 5.3 Accuracy represents usual character accuracy and Accuracy represents equally weighted class accuracy.

$Precision = RL$

$$\frac{Retrieved\ Documents}{Relevant\ Document\ Retrieved} \quad \text{-----}(2)$$

where "*errors*" is a minimum number of edit operations (character insertions, deletions, and substitutions) needed to fully correct the text and "Number_of_characters" is a number of characters in document.

Recall is the second metric. It is the percentage of papers that have been located and are relevant to the query. When a binary classification test correctly detects or excludes a condition, accuracy is used as a statistical measure.

$$Accuracy = \frac{TP+TN+FP+FN}{TP+TN} \quad \text{-------}\ (3)$$

Table 5.3 gives the performance evaluation of OCIR and NBA are analysed by means of existing method like K-means, Support Vector Machine, Naïve Bayes, and proposed methods are compared with proposed automated term based Fuzzy logic analysis for information retrieval (FCA-IR) and OCIR based sentiment analysis for identifying the image into suitable text format.

Table 5.3 Performance Evaluation of OCIR

Methods	**Precision**	**Recall**	**CA**	**Accuracy**
SVM	**90.34**	**90.56**	**91.34**	**91.26**
NBA	**92.45**	**92.67**	**92.41**	**92.71**
FCA-IR	**93.02**	**93.65**	**93.12**	**93.81**
OCIR	**94.69**	**94.39**	**94.26**	**94.11**

Key considerations in the design of a recognition system encompass the delineation of pattern classes, sensing environment, pattern representation, feature extraction and selection, cluster analysis, classifier design and learning, training and test sample selection, and performance evaluation. The fundamental issue of identifying intricate patterns amidst random patterns of varying scale, orientation, and location persists unresolved. Following the classification of the input photographs, they undergo preprocessing tailored to specific classes. The structured representation established in the prior stage facilitates the discovery of association rules, the identification of prevalent keywords, and the execution of sentiment analysis through an OCIR methodology that utilizes a collection of positive and negative terms.

The accuracy of the model in predicting the appropriate stress level for a particular input text can be used to assess the outcomes of stress detection from conversation utilizing natural language processing embeddings and tokenization. Evaluation criteria including precision, recall, and F1 score can be used to gauge the model's performance. The model's capacity to accurately categorize various stress levels can be measured using these criteria. To make sure that the model's performance is consistent across various subsets of the data, cross validation on the dataset can be used to further assess the model's accuracy.

5.2 FUTURE ENHANCEMENTS

The Stress Detection from Chat project can be refined and expanded in the future to increase its precision and efficacy. Adding more sophisticated machine learning techniques, such deep learning or reinforcement learning, to better capture the subtleties of human language and conversation is one possible area for advancement. Furthermore, as there may be a growing need for stress detection in other languages, the project could be expanded to include languages other than English. Investigating the use of wearable technology or other sensors to gather physiological data, such as skin conductance or heart rate, to increase the precision of stress detection is another possible avenue for advancement. To provide a more thorough examination of a person's stress levels, this may entail combining the conversation data with real-time physiological data. Additionally, the initiative can be used in a variety of contexts and industries, including call centers, mental health clinics, and workplaces, to offer prompt interventions and assistance to those who are under a lot of stress.

REFERENCES

[1] Bhamare DP, Suryawanshi P. Review on Reliable Pattern Recognition with Machine Learning Techniques. Fuzzy Information and Engineering, 2018; 10(3): 362–377. https://doi.org/10.1080/16168658.2019.161103.

[2] Dr.S.Brindha, Dr.S.Sukumaran, Enhanced Pattern Classification Methods for Text Categorization Mining, International Journal of All Research Education and Scientific Methods (IJARESM), ISSN: 2455-6211 Volume 10, Issue 10, October-2022, Impact Factor: 7.429, Available online at: www.ijaresm.com.

[3] Armengol E, Boixader D, Grimaldo,F. Special Issue on Pattern Recognition Techniques in Data Mining. Pattern Recognition Letters, 2017; 93: 1–2.
https://doi.org/10.1016/j.patrec.2017.02.014

[4] Ermushev S, Balashov A. A Complex Machine Learning Technique for Ground Target Detection and Classification. Int J Appl Eng Res., 2017; 1(1): 158–161

[5] Omarov B, Cho YI. Machine learning based pattern recognition and classification framework development. In 2017 17th International conference on Control, Automation and Systems (ICCAS 2017). Ramada Plaza, Jeju, Korea, 2017.

[6] Jindal, A., Singh, A., & Kumar, P. (2019). Detecting Stress in Online Chat Conversations using NLP. In 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT) (pp. 1-6). IEEE.

[7] Buechel, S., Hahn, U., & Neal, D. T. (2019). Identifying emotional distress and stress in social media messages. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language

Processing (EMNLP-IJCNLP) (pp. 5308-5317).

[8] Lee, C. H., Lim, Y. M., & Kim, K. J. (2018). A chatbot system for stress management using deep neural network. In Proceedings of the 2018 International Conference on Information and Communication Technology Convergence (ICTC) (pp. 1170-1172). IEEE.

[9] Kudugunta, S., & Rama, T. G. (2018). Detecting stress in email conversations using NLP. In Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers) (pp. 2222-2232).

[10] Khan, A. A., Kumar, M., Srivastava, S., & Bhattacharyya, P. (2020). Detecting stress in online text-based counseling sessions using deep learning. In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP) (pp. 415426).

Authors

Dr.S.Brindha received B.Sc degree in Science from Bharathiyar University. She done her M.Sc in Periyar University and she awarded M.Phil Computer Science from the Bharathiyar University. She received the Ph.D degree in Computer Science from the Bharathiar University. She has 12 years of teaching experience and 6 years of Technical Experience. At present she is working as Assistant Professor in Department of Computer Applications in SRM Faculty of Science and Humanities, SRM Institute of Science and Technology, Kattankulathur, Chengalpattu, Chennai, Tamilnadu, India. She published around 32 research papers in International Journals and Conferences. Published E-learning Concepts and Development tool book in the year of 2021. Published Exploring Taxonomy Based Methods for Detecting Patterns in Text Documents book in the year 2022. In the year 2023, published Data Science and Analytics. In the year 2024. She has published Information Security and Data Privacy. She have Published various book Chapters related to Climate Change and Human Health, Exploring Women Leadership: Achieving an Equal Future, Sustainable Solution for Green Environment. Received Women Researcher Award, Best Young Scientist Award and Best Faculty Award in Research. Member in International Research Awards on Science, Technology and Management. Lifetime Membership in Professional Nobel Sansnow’s Professional Foundation, Approved by Ministry of Corporate Affairs, Government of India. Member in Computer Society of India and Association Computing Machinery (ACM). Her Research area includes Text Mining, Image Processing, Pattern Taxonomy Mining, Deep Learning, Artificial Intelligence and Machine Learning.

Dr. S. Ravichandran is currently working as Head and Professor in the Department of Chemistry at St.Peter’s Institute of Higher Education and

Research, Chennai. He completed his Ph.D. in 2006 from Madurai Kamaraj University, Madurai (Tamilnadu) and M.Sc. from Pondicherry University, Pondicherry. He has qualified in GATE with a score of 95 percentile conducted by Ministry of Human Research and Development in the year 1998. He has 20 years of Teaching and Research experiences and published 154 International papers. He has published 9 patents and 7 Textbooks and 8 book chapters. He has received Bharat Shiksha Ratan award and Lifetime achievement award from Global society in 2012, 2013 from New Delhi. He has also received the award of Academic Excellence by Arab Translators Association, Bahrain on 24th November 2021 in recognition of research publications achievement. Received the Life Time Achievement Award with medal from Blue Bird Welfare Association, Prayagraj in a National Conference on Recent Trends in Science, Technology and Management conducted by Madhu Vachaspati Institute of Engineering and Technology, Kaushambi (UP) on 13th February 2022. Received the Life Time Achievement Award with medal from Sansnow's Nobel Professional Foundation, Kanyakumari, (Tamilnadu) approved by Ministry of Corporate Affairs, Government of India on 4th th June 2022. Very recently he has received the Incredible Researcher of India Award with medal from Record Owner, Government of India, Ahmedabad on 30th August 2022. He has been a Life membership (L36754) in Indian Science Congress Association, Kolkata. His area of research work are: Synthesis of Mannich and Schiff bases with metal ions like Cu(II), Co(II), Ni(II) and Zn (II). Structural characterization using spectroscopic techniques like UV-Vis., IR, 1H NMR, CV, Chromatography and EPR. His current interest is to focus on the development of novel greener methodology for a Sustainable Development.

Sajith Balaji Sarvesh is presently pursuing a degree in Computer Science and Engineering at Karpaga Vinayaga College of Engineering and Technology. He is deeply invested in the intersection of chemistry and computer science engineering. He Published 4 Research Papers in Data Mining. He very interested to join the college event activities and participate in academic roles. His area of interest include human-centric activities, robotics, and the advancement of human and environmental development through innovative technology. Areas of interest include computer games, artificial intelligence, and deep learning.